BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Retrovirus from the HIV group and its use

[001] This application is a continuation of U.S. application No. 09/109,916, filed July 2, 1998, now allowed, which is a divisional application of U.S. application No. 08/468,059, filed June 6, 1995, now U.S. Patent No. 5,840,480, issued November 24, 1998; which is a divisional application of U.S. application No. 08/132,653, filed October 5, 1993, now abandoned; the disclosures of all of which are incorporated herein by reference.

[002] The present invention relates to a novel retrovirus from the HIV group, as well as to variants or parts thereof which contain the essential properties of the virus. A process is described for culturing the retrovirus. The invention furthermore relates to the isolation of this retrovirus and to use of the virus, its parts or extracts for medicinal purposes, for diagnostics and in the preparation of vaccines.

[003] Retroviruses which belong to the so-called HIV group lead in humans who are infected by them to disease manifestations which are summarized under the collective term immunodeficiency or AIDS (acquired immune deficiency syndrome).

[004] Epidemiological studies verify that the human immunodeficiency virus (HIV) represents the etiological agent in the vast majority of AIDS (acquired immune deficiency syndrome) cases. A retrovirus which was isolated from a patient and characterized in 1983 received the designation

HIV-1 (Barré-Sinoussi, F. et al., Science 220, 868-871 [1983]). A variant of HIV-1 is described in WO 86/02383.

[005] A second group of human immunodeficiency viruses was identified in 1985 in West Africa (Clavel, F. et al., Science 233, 343-346 [1986]) and designated human immunodeficiency virus type 2 (HIV-2) (EP-A-0 239 425). While HIV-2 retroviruses clearly differ from HIV-1, they do exhibit affinity with simian immunodeficiency viruses (SIV-2). Like HIV-1, HIV-2 also leads to AIDS symptomatology.

[006] A further variant of an immunodeficiency retrovirus is described in EP-A-0 345 375 and designated there as HIV-3 retrovirus (ANT 70).

[007] The isolation of a further, variant, immunodeficiency virus is also described in Lancet Vol. 340, Sept. 1992, pp. 681-682.

[008] It is characteristic of human immunodeficiency viruses that they exhibit a high degree of variability, which significantly complicates the comparability of the different isolates. For example, when diverse HIV-1 isolates are compared, high degrees of variability are found in some regions of the genome while other regions are comparatively well conserved (Benn, S. et al., Science 230, 949-951 (1985]). It was also possible to observe an appreciably greater degree of polymorphism in the case of HIV-2 (Clavel, F. et al., Nature 324, 691-695 [1986]). The greatest degree of genetic stability is possessed by regions in the gag and pol genes which encode proteins which are essential for structural and

enzymic purposes; some regions in the env gene, and the genes (vif, vpr, tat, rev and nef) encoding regulatory proteins, exhibit a high degree of variability. In addition to this, it was possible to demonstrate that antisera against HIV-1 also crossreact with gag and pol gene products from HIV-2 even though there was only a small degree of sequence homology. Little hybridization of significance likewise took place between these two viruses unless conditions of very low stringency were used (Clavel, F. et al., Nature 324, 691-695 [1986]).

[009] Owing to the wide distribution of retroviruses from the HIV group and to the fact that a period of a few to many years (2-20) exists between the time of infection and the time at which unambiguous symptoms of pathological changes are recognizable, it is of great importance from the epidemiological point of view to determine infection with retroviruses of the HIV group at as early a stage as possible and, above all, in a reliable manner. This is not only of importance when diagnosing patients who exhibit signs of immunodeficiency, but also when monitoring blood donors. It has emerged that, when retroviruses of the HIV-1 or HIV-2 type, or components thereof, are used in detection systems, antibodies can either not be detected or only detected weakly in many sera even though signs of immunodeficiency are present in the patients from which the sera are derived. In certain cases, such detection is possible using the retrovirus from the HIV group according to the invention.

[010] This patent describes the isolation and characterization of a novel human immunodeficiency virus, designated below as MVP-5180/91 (SEQ ID NO:56), which was isolated from the peripheral lymphocytes of a female patient from the Cameroons who was 34 years old in 1991 and who exhibited signs of immunodeficiency. From the point of view of geography, this retrovirus originates from a region in Africa which is located between West Africa, where there is endemic infection with HIV-2 and HIV-1 viruses, and Eastern Central Africa, where it is almost exclusively HIV-1 which is disseminated. Consequently, the present invention relates to a novel retrovirus, designated MVP-5180/91 (SEQ ID NO:56), of the HIV group and its variants, to DNA sequences, amino acid sequences and constituent sequences derived therefrom, and to test kits containing the latter. The retrovirus MVP-5180/91 (SEQ ID NO:56) has been deposited with the European Collection of Animal Cell Cultures (ECACC) PHLS Centre for Applied Microbiology & Research, Porton Down, Salisbury Wilts. SP4 0JG, United Kingdom, on September 23, 1992 under ECACC Accession No. V 920 92 318 in accordance with the stipulations of the Budapest Treaty. The ECACC is located at the PHLS Centre for Applied Microbiology & Research, Porton Down, Salisbury, Wilts, SP4 0JG, U.K. The deposit was made on September 23, 1992, and was assigned Accession No. V 920 92 318. The date of notification of acceptance of the culture was January 21, 1993.

[011] As do HIV-1 and HIV-2, MVP-5180/91 (SEQ ID NO:56) according to the invention grows in the following cell lines:

HUT 78, Jurkat cells, C8166 cells and MT-2 cells. The isolation and propagation of viruses is described in detail in the book "Viral Quantitation in HIV Infection, Editor Jean-Marie Andrieu, John Libbey Eurotext, 1991". The procedural methods described in that publication are by reference made a subject of the disclosure of the present application.

[012] In addition to this, the virus according to the invention possesses a reverse transcriptase which is magnesium-dependent but not manganese-dependent. This represents a further property possessed in common with the HIV-1 and HIV-2 viruses.

[013] In order to provide a better understanding of the differences between the MVP-5180/91 (SEQ ID NO:56) virus according to the invention and the HIV-1 and HIV-2 retroviruses, the construction of the retroviruses which cause immunodeficiency will first of all be explained in brief. Within the virus, the RNA is located in a conical core which is assembled from protein subunits which carry the designation p 24 (p for protein). This inner core is surrounded by a protein coat, which is constructed from protein p 17 (outer core), and by a glycoprotein coat which, in addition to lipids, which originate from the host cell, contains the transmembrane protein gp 41 and the coat protein 120 (gp 120). This gp 120 can then bind to the CD-4 receptors of the host cells.

[014] Brief Description of the Drawings

[015] Figure 1 depicts the arrangement of the genome of retroviruses of the HIV type.

[016] Figure 2 is a graph depicting the binding affinity for the monoclonal antibody p24 in relation to the content of reverse transcriptase for the retroviruses HIV-1, HIV-2, and MVP-5180/91.

[017] Figure 3 depicts a western blot of MVP-5180/91 and HIV-1, isolated from German patients.

[018] Figure 4 depicts the almost complete nucleic acid sequence of the retrovirus MVP-5180/91.

[019] Figure 5 depicts the strategy for PCR amplification, cloning, and sequencing of MVP-5180/91.

[020] Figure 6 depicts a comparison of the sequence in Figure 4 and the sequence obtained using the PCR amplification techniques depicted in Figure 5.

[021] Figure 7 depicts a comparison of the amino acid sequences of the GAG protein determined from the sequence of Figure 4 with the GAG protein sequence obtained using the PCR amplification techniques depicted in Figure 5.

[022] Figure 8 depicts the immunological specificities of the V3 loop of HIV-1, HIV-2, and MVP-5180/91.

[023] Detailed Description of the Invention

[024] As far as is known, the RNA of HIV viruses - portrayed in a simplified manner - possesses the following gene regions: so-called long terminal repeats (LTR) at each end, together with the following gene regions: gag, pol, env and nef. The gag gene encodes, inter alia, the core proteins, p 24 and p 17, the pol gene encodes, inter alia, the reverse transcriptase, the RNAse H and the integrase, while the env gene encodes the gp 41 and gp 120 glycoproteins of the virus

coat. The nef gene encodes a protein having a regulatory function. The arrangement of the genome of retroviruses of the HIV type is shown diagrammatically in Figure 1.

[025] The HIV-1 and HIV-2 retroviruses can be distinguished, inter alia, by testing viral antigen using a monoclonal antibody which is commercially available from Abbott (HIVAG-1 monoclonal) in the form of a test kit and is directed against (HIV-1) p 24. It is known that the content of reverse transcriptase is roughly the same in the HIV-1 and HIV-2 virus types. If, therefore, the extinction (E 490 nm.) obtained in dilutions of the disrupted viruses by means of the antigen-antibody reaction is plotted against the activity of reverse transcriptase, a series of graphs is obtained corresponding roughly to that in Figure 2. In this context, it is observed that, in the case of HIV-1, the monoclonal antibody employed has a very high binding affinity for p 24 in relation to the content of reverse transcriptase. By contrast, the monoclonal antibody employed has only a very low binding affinity for p 24 in the case of HIV-2, once again in relation to the content of reverse transcriptase. If these measurements are carried out on MVP-5180/91 (SEQ ID NO:56), the curve is then located almost precisely in the centre between the curves for HIV-1 and HIV-2, i.e. the binding affinity of the monoclonal antibody for MVP-5180/91 p 24 is reduced as compared with the case of HIV-1. Figure 2 shows this relationship diagrammatically, with RT denoting reverse transcriptase, and the protein p 24, against which is directed the monoclonal antibody which is present in the test kit which

can be purchased from Abbott, being employed as the antigen (Ag).

[026] The so-called PCR (polymerase chain reaction) system has proved to have a multiplicity of uses in genetic manipulation, and the components which are required for implementing the process can be purchased. Using this process, it is possible to amplify DNA sequences if regions of the sequence to be amplified are known. Short, complementary DNA fragments (oligonucleotides = primers) have then to be synthesized which anneal to a short region of the nucleic acid sequence to be amplified. For carrying out the test, HIV nucleic acids are introduced together with the primers into a reaction mixture which additionally contains a polymerase and nucleotide triphosphates. The polymerization (DNA synthesis) is carried out for a given time and the nucleic acid strands are then separated by heating. After cooling, the polymerization then proceeds once more. If, therefore, the retrovirus according to the invention is an HIV-1 or HIV-2 virus, it should be possible to amplify the nucleic acid using primers which are conserved within the known sequences of the HIV-1 and HIV-2 viruses. Some primers of this type have previously been described (Lauré, F. et al., Lancet ii, (1988) 538-541 for pol 3 and pol 4, and Ou C.Y. et al., Science 239 (1988) 295-297 for sk 38/39 and sk 68/69).

[027] It was discovered that use of particular primer pairs having the following sequence:

gaga (SEQ ID NO:1): CTACT AGTAC CCTTC AGG

gagb (SEQ ID NO:2): CGGTC TACAT AGTCT CTAAA G

sk38 (SEQ ID NO:3): CCACC TATCC CAGTA GGAGA A

sk39 (SEQ ID NO:4): CCTTT GGTCC TTGTC TTATG TCCAG AATGC

or

pol3 (SEQ ID NO:5): TGGGA AGTTC AATTA GGAAT ACCAC

pol4 (SEQ ID NO:6): CCTAC ATAGA AATCA TCCAT GTATT G

pol3n (SEQ ID NO:7): TGGAT GTGGG TGATG CATA

pol4n (SEQ ID NO:8): AGCAC ATTGT ACTGA TATCT A

and

SK145 (SEQ ID NO:9): AGTGG GGGGA CATCA AGCAG CC

SK150 (SEQ ID NO:10): TGCTA TGTCA CTTCC CCTTG GT

145-P (SEQ ID NO:11): CCATG CAAAT GTTAA AAGAG AC

150-P (SEQ ID NO:12): GGCCT GGTGC AATAG GCCC

or a combination of pol 3 and pol 4 with

UNI-1 (SEQ ID NO:13): GTGCT TCCAC AGGGA TGGAA

UNI-2 (SEQ ID NO:14): ATCAT CCATG TATTG ATA

(Donehower L.A. et al. (1990) J. Virol. Methods 28, 33-46) and employing PCR with nested primers, led to weak amplifications of the MVP-5180/91 DNA (SEQ ID NO:56).

[028] No amplification, or only weak amplification as compared with HIV-1, possibly attributable to impurities, was obtained with the following primer sequences:

tat 1 (SEQ ID NO:15): AATGG AGCCA GTAGA TCCTA

tat 2 (SEQ ID NO:16): TGTCT CCGCT TCTTC CTGCC

tat 1P (SEQ ID NO:17): GAGCC CTGGA AGCAT CCAGG

tat 2P (SEQ ID NO:18): GGAGA TGCCT AAGGC TTTTG

enva (SEQ ID NO:19): TGTTC CTTGG GTTCT TG

envb (SEQ ID NO:20): GAGTT TTCCA GAGCA ACCCC

sk68 (SEQ ID NO:21): AGCAG CAGGA AGCAC TATGG

sk69 (SEQ ID NO:22): GCCCC AGACT GTGAG TTGCA ACAG

5v3e (SEQ ID NO:23): GCACA GTACA ATGTA CACAT GG

3v3e (SEQ ID NO:24): CAGTA GAAAA ATTCC CCTCC AC

5v3degi (SEQ ID NO:25): TCAGG ATCCA TGGGC AGTCT AGCAG AAGAA G

3v3degi (SEQ ID NO:26): ATGCT CGAGA ACTGC AGCAT CGATT CTGGG TCCCC TCCTG AG

3v3longdegi (SEQ ID NO:27): CGAGA ACTGC AGCAT CGATG CTGCT CCCAA GAACC CAAGG

3v3longext (SEQ ID NO:28): GGAGC TGCTT GATGC CCCAG A

gagdi (SEQ ID NO:29): TGATG ACAGC ATGTC AGGGA GT

pol e (SEQ ID NO:30): GCTGA CATTT ATCAC AGCTG GCTAC

[029] Amplifications which were weak as compared with those for HIV-1, but nevertheless of the same intensity as those for the HIV-2 isolate (MVP-11971/87) employed, were obtained with

gag c (SEQ ID NO:31): TATCA CCTAG AACTT TAAAT GCATG GG

gag d (SEQ ID NO:32): AGTCC CTGAC ATGCT GTCAT CA

env c (SEQ ID NO:33): GTGGA GGGGA ATTTT TCTAC TG

env d (SEQ ID NO:34): CCTGC TGCTC CCAAG AACCC AAGG.

[030] The so-called Western blot (immunoblot) is a common method for detecting HIV antibodies. In this method, the viral proteins are fractionated by gel electrophoresis and then transferred to a membrane. The membranes provided with the transferred proteins are then brought into contact with sera from the patients to be investigated. If antibodies against the viral proteins are present, these antibodies will bind to the proteins. After the membranes have been washed, only antibodies which are specific for the viral proteins will remain. The antibodies are then rendered visible using antiantibodies which, as a rule, are coupled to an enzyme which catalyzes a color reaction. In this way, the bands of the viral proteins can be rendered visible.

[031] The virus MVP-5180/91 (SEQ ID NO:56) according to the invention exhibits two significant and important differences from the HIV-1 and HIV-2 viruses in a Western blot. HIV-1 regularly shows a strong band, which is attributable to protein p 24, and a very weak band, which is often scarcely visible and which is attributable to protein p 23. HIV-2 exhibits a strong band, which is attributable to protein p 25, and sometimes a weak band, which is attributable to protein p 23. In contrast to this, the MVP-5180/91 (SEQ ID NO:56) virus according to the invention exhibits two bands of

approximately equal strength, corresponding to proteins p 24 and p 25.

[032] A further significant difference exists in the bands which are attributable to reverse transcriptase. HIV-1 shows one band (p 53) which corresponds to reverse transcriptase and one band (p 66) which corresponds to reverse transcriptase bound to RNAse H. In the case of HIV-2, the reverse transcriptase corresponds to protein p 55 and, if it is bound to RNAse H, to protein p 68. By contrast, MPV-5180/91 (SEQ ID NO:56) according to the invention exhibits one band at protein p 48, which corresponds to reverse transcriptase, and one band, at protein p 60, which corresponds to reverse transcriptase bound to RNAse H. It can be deduced from these results that the reverse transcriptase of MVP-5180/91 (SEQ ID NO:56) has a molecular weight which is roughly between 3 and 7 kilodaltons less than that of the reverse transcriptases of HIV-1 and HIV-2. The reverse transcriptase of MVP-5180 consequently has a molecular weight which is roughly between 4,500 daltons and 5,500 daltons less than that of the reverse transcriptase of HIV-1 or HIV-2.

[033] It was discovered that anti-env antibodies could only be detected weakly in the sera of German patients exhibiting signs of immunodeficiency when the MVP-5180/91 (SEQ ID NO:56) virus according to the invention was used, whereas the sera reacted strongly if an HIV-1 virus was used instead of the virus according to the invention. This stronger detection reaction was located in the gp 41 protein, in particular. In the experiments, serum panels were compared

which on the one hand derived from German patients and on the other from African patients showing signs of immune deficiency.

[034] The abovementioned characteristics are indicative of those virus variants which correspond to MVP-5180/91 (SEQ ID NO:56) according to the invention. Therefore, the virus according to the invention, or variants thereof, can be obtained by isolating immunodeficiency viruses from heparinized donor blood derived from persons who exhibit signs of immune deficiency and who preferably originate from Africa.

[035] Since the virus possessing the abovementioned properties has been isolated, the cloning of a cDNA can be carried out in the following manner: the virus is precipitated from an appropriately large quantity of culture (about 1 l) and then taken up in phosphate-buffered sodium chloride solution. It is then pelleted through a (20% strength) sucrose cushion. The virus pellet can be suspended in 6 M guanidinium chloride in 20 mM dithiothreitol and 0.5% Nonidet P 40. CsCl is added to bring its concentration to 2 molar and the solution containing the disrupted virus is transferred to a cesium chloride cushion. The viral RNA is then pelleted by centrifugation, and subsequently dissolved, extracted with phenol and precipitated with ethanol and lithium chloride. Synthesis of the first cDNA strand is carried out on the viral RNA, or parts thereof, using an oligo(dT) primer. The synthesis can be carried out using a commercially available kit and adding reverse transcriptase. To synthesize the second strand, the RNA strand of the RNA/DNA

hybrid is digested with RNase H, and the second strand is then synthesized using E. coli DNA polymerase I. Blunt ends can then be produced using T4 DNA polymerase and these ends can be joined to suitable linkers for restriction cleavage sites. Following restriction digestion with the appropriate restriction endonuclease, the cDNA fragment is isolated from an agarose gel and ligated to a vector which has previously been cut in an appropriate manner. The vector containing the cDNA insert can then be used for transforming competent E. coli cells. The colonies which are obtained are then transferred to membranes, lysed and denatured, and then finally detected by hybridization with nucleic acid labeled with digoxigenin or biotin. Once the corresponding cDNA has been prepared by genetic manipulation, it is possible to isolate the desired DNA fragments originating from the retrovirus. By incorporating these fragments into suitable expression vectors, the desired protein or protein fragment can then be expressed and employed for the diagnostic tests.

[036] As an alternative to the stated method, the immunodeficiency virus can be cloned with the aid of PCR technology, it being possible to use the abovementioned primers.

[037] The similarity between different virus isolates can be expressed by the degree of homology between the nucleic acid or protein sequences. 50% homology means, for example, that 50 out of 100 nucleotides or amino acid positions in the sequences correspond to each other. The homology of proteins

is determined by sequence analysis. Homologous DNA sequences can also be identified by the hybridization technique.

[038] In accordance with the invention, a part of the coat protein was initially sequenced and it was ascertained that this sequence possessed only relatively slight homology to the corresponding sequences from viruses of the HIV type. On the basis of a comparison with HIV sequences, which was carried out using data banks, it was established, in relation to the gp 41 region in particular, that the homology was at most 66% (nucleotide sequence).

[039] In addition to this, the region was sequenced which encodes gp 41. This sequence is presented in Tables 1 and 3. Table 1 includes DNA SEQ ID NO:37, DNA SEQ ID NO:38, and amino acid SEQ ID NO:39. Table 3 includes DNA SEQ ID NO:44, DNA SEQ ID NO:45, and amino acid SEQ ID NO:46.

[040] The present invention therefore relates to those viruses which possess an homology of more than 66%, preferably 75% and particularly preferably 85%, to the HIV virus, MVP-5180/91 (SEQ ID NO:56), according to the invention, based on the nucleotide sequence in Table 1 (SEQ ID NO:37; SEQ ID NO:38) and/or in Table 3 (SEQ ID NO:44; SEQ ID NO:45).

[041] Furthermore, the present invention relates to those viruses which possess an homology of more than 66%, preferably 75% and particularly preferably 85%, to partial sequences of the nucleotide sequence presented in Table 3 (SEQ ID NO:44; SEQ ID NO:45), which sequences are at least 50, preferably 100, nucleotides long. This corresponds to a

length of the peptides of at least 16, and preferably of at least 33, amino acids.

[042] The sequence of the virus according to the invention differs from that of previously known viruses. The present invention therefore relates to those viruses, and corresponding DNA and amino acid sequences, which correspond to a large extent to the sequence of the virus according to the invention, the degree of deviation being established by the degree of homology. An homology of, for example, more than 85% denotes, therefore, that those sequences are included which have in at least 85 of 100 nucleotides or amino acids the same nucleotides or amino acids, respectively, while the remainder can be different. When establishing homology, the two sequences are compared in such a way that the greatest possible number of nucleotides or amino acids corresponding to each other are placed in congruence.

[043] The (almost) complete sequence, given as the DNA sequence of the virus according to the invention, is reproduced in Fig. 4 and included as DNA SEQ ID NO:56. In this context, the present invention relates to viruses which possess the sequence according to Fig. 4 (SEQ ID NO:56), and variants thereof which possess a high degree of homology with the sequence of Fig. 4 (SEQ ID NO:56), as well as proteins, polypeptides and oligopeptides derived therefrom which can be used diagnostically or can be employed as vaccines.

[044] Using the isolated sequence as a basis, immunodominant epitopes (peptides) can be designed and synthesized. Since the nucleic acid sequence of the virus is

known, the person skilled in the art can derive the amino acid sequence from this known sequence. A constituent region of the amino acid sequence is given in Table 3 (SEQ ID NO:46). The present invention also relates, therefore, to antigens, i.e. proteins, oligopeptides or polypeptides, which can be prepared with the aid of the information disclosed in Figure 4 (SEQ ID NO:56) and Table 3 (SEQ ID NO:44; SEQ ID NO:45, and SEQ ID NO:46). These antigens, proteins, polypeptides and oligopeptides possess amino acid sequences which can either be derived from Figure 4 (SEQ ID NO:56) or are given in Table 3 (SEQ ID NO:46). The antigens or peptides can possess relatively short constituent sequences of an amino acid sequence which is reproduced in Table 3 (SEQ ID NO:46) or which can be derived from Figure 4 (SEQ ID NO:56). This amino acid sequence is at least 6, preferably at least 10 and particularly preferably at least 15, amino acids in length. These peptides can be prepared not only with the aid of recombinant technology but also using synthetic methods. A suitable preparation route is solid-phase synthesis of the Merrifield type. Further description of this technique, and of other processes known to the state of the art, can be found in the literature, e.g. M. Bodansky, et al., Peptide Synthesis, John Wiley & Sons, 2nd Edition 1976.

[045] In the diagnostic tests, a serum sample from the person to be investigated is brought into contact with the protein chains of one or more proteins or glycoproteins (which can be expressed in eukaryotic cell lines), or parts thereof, which originate from MVP-5180/91 (SEQ ID NO:56). Test

processes which are preferred include immunofluorescence or immunoenzymatic test processes (e.g. ELISA or immunoblot).

[046] In the immunoenzymatic tests (ELISA), antigen originating from MVP-5180/91 (SEQ ID NO:56) or a variant thereof, for example, can be bound to the walls of microtiter plates. The dosage used in this context depends to an important degree on the test system and the treatment of the microtiter plates. Serum or dilutions of serum deriving from the person to be investigated are then added to the wells of the microtiter plates. After a predetermined incubation time, the plate is washed and specific immunocomplexes are detected by antibodies which bind specifically to human immunoglobulins and which had previously been linked to an enzyme, for example horseradish peroxidase, alkaline phosphatase, etc., or to enzyme-labeled antigen. These enzymes are able to convert a colorless substrate into a strongly colored product, and the presence of specific anti-HIV antibodies can be gathered from the strength of the coloration. A further option for using the virus according to the invention in test systems is its use in Western blots.

[047] Even if the preparation of vaccines against immunodeficiency diseases is proving to be extremely difficult, this virus, too, or parts thereof, i.e. immunodominant epitopes and inducers of cellular immunity, or antigens prepared by genetic manipulation, can still be used for developing and preparing vaccines.

[048] **Example 1:**

[049] The immunodeficiency virus according to the invention, MVP-5180/91 (SEQ ID NO:56), was isolated from the blood of a female patient exhibiting signs of immune deficiency. To do this, peripheral mononuclear cells (peripheral blood lymphocytes, PBL) and peripheral lymphocytes from the blood (PBL) of a donor who was not infected with HIV were stimulated with phytohemagglutinin and maintained in culture. For this purpose, use was made of the customary medium RPMI 1640 containing 10% fetal calf serum. The culture conditions are described in Landay A. et al., J. Inf. Dis., 161 (1990) pp. 706-710. The formation of giant cells was then observed under the microscope. The production of HIV viruses was ascertained by determining the p 24 antigen using the test which can be purchased from Abbott. An additional test for determining the growth of the viruses consisted of the test using particle-bound reverse transcriptase (Eberle J., Seibl R., J. Virol. Methods 40, 1992, pp. 347-356). The growth of the viruses was therefore determined once or twice a week on the basis of the enzymatic activities in the culture supernatant, in order to monitor virus production. New donor lymphocytes were added once a week.

[050] Once it was possible to observe HIV virus multiplication, fresh peripheral lymphocytes from the blood (PBL) of healthy donors, who were not infected with HIV, were infected with supernatant from the first culture. This step was repeated and the supernatant was then used to infect H 9 and HUT 78 cells. In this way, it was possible to achieve

permanent production of the immunodeficiency virus. The virus was deposited with the ECACC under No. V 920 92 318.

[051] **Example 2:**

[052] So-called Western blot or immunoblot is currently a standard method for detecting HIV infections. Various sera were examined in accordance with the procedure described by Gürtler et al. in J. Virol. Meth. 15 (1987) pp. 11-23. In doing this, sera from German patients were compared with sera which had been obtained from African patients. The following results were obtained:

Virus type	German sera	African sera
HIV-1, virus isolated from German patients	strong reaction	strong reaction using gp 41
MVP-5180/91 (SEQ ID NO:56)	no reaction to weak reaction using gp 41	strong reaction

[053] The results presented above demonstrate that a virus of the HIV-1 type isolated from German patients may possibly, if used for detecting HIV infections, fail to provide unambiguous results if the patient was infected with a virus corresponding to MVP-5180/91 (SEQ ID NO:56) according to the invention. It is assumed here that those viruses can be detected using the virus according to the invention which possess at least about 85% homology, based on the total genome, with the virus according to the invention.

[054] **Example 3:**

[055] Further Western blots were carried out in accordance with the procedure indicated in Example 2. The results are presented in the enclosed Figure 3. In this test, the viral protein of the immunodeficiency virus MVP-5180/91 (SEQ ID NO:56) according to the invention, in the one case, and the viral protein of an HIV-1 type virus (MVP-899), in the other, was fractionated by gel electrophoresis and then transferred to cellulose filters. These filter strips were incubated with the sera from different patients and the specific antibodies were then rendered visible by a color reaction. The left half of the figure with the heading MVP-5180 shows the immunodeficiency virus according to the invention. The right half of the figure shows a virus (MVP-899), which is an HIV-1 virus, isolated from a German donor.

[056] In Figure 3, the same sera (from German patients) were in each case reacted with two respective filter strips, the numbers 8 and 26; 9 and 27; 10 and 28; 11 and 29; 12 and 30; 13 and 31; 14 and 32; 15 and 33, and 16 and 34 indicating the same sera. Sera from African patients were employed in the Western blots having the numbers 17 and 18. The numbers on the right hand margins indicate the approximate molecular weights in thousands (KD).

[057] Figure 3 shows clearly that sera from German patients only react very weakly with the immunodeficiency virus according to the invention in a Western blot using gp 41. By contrast, sera from African patients react very strongly with the immunodeficiency virus according to the

invention. Figure 3 makes it clear, therefore, that when the immunodeficiency virus according to the invention is used those immunodeficiency infections can be detected which only yield questionable, i.e. not unambiguously positive, results when an HIV-1 or HIV-2 virus is used. This option for detection can be of far-reaching diagnostic importance since, in those cases in which only questionable results are obtained in a Western blot, it cannot be established with unambiguous certainty whether an infection with an immunodeficiency virus is present. However, if the immunodeficiency virus according to the invention can be used to assign such questionable results to an infection with a virus of the type according to the invention, this then represents a substantial diagnostic advance.

[058] **Example 4:**

[059] DNA isolation, amplification and structural characterization of sections of the genome of the HIV isolate MVP-5180/91 (SEQ ID NO:56).

[060] Genomic DNA from HUT 78 cells infected with MVP-5180/91 (SEQ ID NO:56) was isolated by standard methods.

[061] In order to characterize regions of the genome of the isolate MVP-5180/91 (SEQ ID NO:56), PCR (polymerase chain reaction) experiments were carried out using a primer pair from the region of the coat protein gp 41. The PCR experiments were carried out in accordance with the method of Saiki et al. (Saiki et al., Science 239: 487-491, 1988) using the following modifications: for the amplification of regions of HIV-specific DNA, 5 µl of genomic DNA from HUT 78 cells

infected with MVP-5180/91 (SEQ ID NO:56) were pipetted into a 100 µl reaction mixture (0.25 mM dNTP, in each case 1 µm primer 1 and primer 2, 10 mM Tris HCl, pH 8.3, 50 mM KCl, 1.5 $MgCl_2$, 0.001% gelatin, 2.5 units of Taq polymerase (Perkin Elmer)), and amplification was then carried out in accordance with the following temperature program: 1. initial denaturation: 3' 95°C, 2. amplification: 90" 94°C, 60" 56°C, 90", 72°C (30 cycles).

[062] The primers used for the PCR and for nucleotide sequencing were synthesized on a Biosearch 8750 oligonucleotide synthesizer.

Primer 1 (SEQ ID NO:35): AGC AGC AGG AAG CAC TAT GG (coordinates from HIV-1 isolate HXB2: bases 7795-7814, corresponds to primer sk 68) (SEQ ID NO:21)

Primer 2 (SEQ ID NO:36): GAG TTT TCC AGA GCA ACC CC (coordinates from HIV-1 isolate HXB2: bases 8003-8022, corresponds to primer env b (SEQ ID NO:20).

[063] The amplified DNA was fractionated on a 3% "Nusieve" agarose gel (from Biozyme) and the amplified fragment was then cut out and an equal volume of buffer (1 * TBE (0.09 M Tris borate, 0.002 M EDTA, pH 8.0) was added to it. After incubating the DNA/agarose mixture at 70°C for 10 minutes, and subsequently extracting with phenol, the DNA was precipitated from the aqueous phase by adding 1/10 vol of 3 M NaAc, pH 5.5, and 2 vol of ethanol and storing at -20°C for 15', and then subsequently pelleted in a centrifuge (Eppendorf) (13,000 rpm, 10', 4°C). The pelleted DNA was dried and taken up in water, and then, after photometric

determination of the DNA concentration at 260 nm in a spectrophotometer (Beckman), sequenced by the Sanger method (F. Sanger, Proc. Natl. Acad,. Sci., 74: 5463, 1977). Instead of sequencing with Klenow DNA polymerase, the sequencing reaction was carried out using a kit from Applied Biosystems ("Taq dye deoxy terminator cycle sequencing", order No.: 401150). Primer 1 (SEQ ID NO:35) or primer 2 (SEQ ID NO:36) (in each case 1 μM) was employed as primers in separate sequencing reactions. The sequencing reaction was analysed on a 373A DNA sequencing apparatus (Applied Biosystems) in accordance with the instructions of the apparatus manufacturer.

[064] The nucleotide sequence of the amplified DNA region, and the amino acid sequence deduced from it, are presented in Table 1. Table 1 includes the DNA sequences SEQ ID NO:37 and SEQ ID NO:38, as well as amino acid SEQ ID NO:39. The top line in Table 1 corresponds to SEQ ID NO:37, the middle line corresponds to SEQ ID NO:38, and the bottom line corresponds to the amino acid SEQ ID NO:39.

[065] Table 1:

```
GCGCAGCGGCAACAGCGCTGACGGTACGGACCCACAGTGTACTGAAGGGTATAGTGCAAC
----------+-------+---------+---------+---------+---------+---------+
CGCGTCGCCGTTGTCGCGACTGCCATGCCTGGGTGTCACATGACTTCCCATATCACGTTG
  A   A   A   T   A   L   T   V   R   T   H   S   V   L   K   G   I   V   Q   Q

AGCAGGACAACCTGCTGAGAGCGATACAGGCCCAGCAACACTTGCTGAGGTTATCTGTAT
--------+--------+---------+---------+----+----+---------+---------+
TCGTCCTGTTGGACGACTCTCGCTATGTCCGGGTCGTTGTGAACGACTCCAATAGACATA
  Q   D   N   L   L   R   A   I   Q   A   Q   Q   H   L   L   R   L   S   V   W
```

```
GGGGTATTAGACAACTCCGAGCTCGCCTGCAAGCCTTAGAAACCCTTATACAGAATCAGC
---------+--------+----------+---------+---------+---------+
CCCCATAATCTGTTGAGGCTCGAGCGGACGTTCGGAATCTTTGGGAATATGTCTTAGTCG

  G   I   R   Q   L   R   A   R   L   Q   A   L   E   T   L   I   Q   N   Q   Q

AACGCCTAAACCTAT
---------+-----  195
TTGCGGATTTGGATA

  R   L   N   L
```

[066] **Example 5:**

[067] The found nucleotide sequence from Table 1 was examined for homologous sequences in the GENEBANK database (Release 72, June 1992) using the GCG computer program (Genetic Computer Group, Inc., Wisconsin USA, Version 7.1, March 1992). Most of the nucleotide sequences of immunodeficient viruses of human origin and of isolates from primates known by July 1992 are contained in this database.

[068] The highest homology shown by the nucleotide sequence from Table 1, of 66%, is to a chimpanzee isolate. The highest homology shown by the investigated DNA sequence from MVP-5180/91 (SEQ ID NO:56) to HIV-1 isolates is 64%. The DNA from Table 1 is 56% homologous to HIV-2 isolates. Apart from the chimpanzee isolate sequence, the best homology between the nucleotide sequence from Table 1 (SEQ ID NO:37; SEQ ID NO:38) and segments of DNA from primate isolates (SIV: simian immunodeficiency virus) is found with a DNA sequence encoding a part of the coat protein region from the SIV isolate (African long-tailed monkey) TYO-1. The homology is 61.5%.

[069] **Example 6:**

[070] The found amino acid sequence from Table 1 (SEQ ID NO:39) was examined for homologous sequences in the SWISSPROT protein database (Release 22, June 1992) using the GCG computer program. Most of the protein sequences of immunodeficiency viruses of human origin and of isolates from primates known by June 1992 are contained in this database.

[071] The highest homology shown by the amino acid sequence from Table 1 (SEQ ID NO:39), of 62.5%, is to a segment of coat protein from the abovementioned chimpanzee isolate. The best homology among HIV-1 coat proteins to the amino acid sequence from Table 1 (SEQ ID NO:39) is found in the isolate HIV-1 Mal. The homology is 59%. The highest homology of the amino acid sequence from Table 1 (SEQ ID NO:39) to HIV-2 coat proteins is 52% (isolate HIV-2 Rod). Since HIV-1 and HIV-2 isolates, themselves, are at most only 64% identical in the corresponding protein segment, the MVP-5180/91 (SEQ ID NO:56) isolate appears to be an HIV variant which clearly differs structurally from HIV-1 and HIV-2 and thus represents an example of an independent group of HIV viruses.

[072] The amino acid sequence of the amplified region of DNA (Table 1; SEQ ID NO:39) from the HIV isolate MVP-5180/91 (SEQ ID NO:56) overlaps an immunodiagnostically important region of the coat protein gp 41 from HIV-1 (amino acids 584-618*) (Table 2, which includes SEQ ID NO:61 as the top line and SEQ ID NO:63 as the bottom line) (Gnann et al., J. Inf.

Dis. 156: 261-267, 1987; Norrby et al., Nature, 329: 248-250, 1987).

[073] Corresponding amino acid regions from the coat proteins of HIV-2 and SIV are likewise immunodiagnostically conserved (Gnann et al., Science, pp. 1346-1349, 1987). Thus, peptides from this coat protein region of HIV-1 and HIV-2 are employed as solid-phase antigens in many commercially available HIV-1/2 antibody screening tests. Approximately 99% of the anti-HIV-1 and anti-HIV-2 positive sera can be identified by them.

[074] The amino acid region of the MVP-5180/91 coat protein (Table 1) could be of serodiagnostic importance owing to the overlap with the immunodiagnostically important region from gp 41. This would be the case particularly if antisera from HIV-infected patients failed to react positively with any of the commercially available antibody screening tests. In these cases, the infection could be with a virus which was closely related to MVP-5180/91 (SEQ ID NO:56).

[075] Table 2: (includes SEQ ID NO:61 as the top line and SEQ ID NO:63 as the bottom line)

```
........RILAVERYLKDQQLLGIWGCSGKLICTTAVPWNAS
        |: |:| .:.:|| |.:
WGIRQLRARLQALETLIQNQQRLNL.................
```

[076] **Example 7:**

[077] DNA isolation, amplification and structural characterization of genome segments from the HIV isolate MVP-5180/91 (SEQ ID NO:56) (encoding gp 41).

[078] Genomic DNA from MVP-5180/91 (SEQ ID NO:56)-infected HUT 78 cells was isolated as described.

[079] In order to characterize genomic regions of the isolate MVP-5180/91, PCR (polymerase chain reaction) experiments were carried out using primer pairs from the gp 41 coat protein region. PCR (Saiki et al., Science 239: 487-491, 1988) and inverse PCR (Triglia et al., Nucl. Acids, Res. 16: 8186, 1988) were carried out with the following modifications:

[080] 1. PCR

[081] For the amplification of HIV-specific DNA regions, 5 µl (218 µg/ml) of genomic DNA from MVP-5180/91-infected HUT 78 cells were pipetted into a 100 µl reaction mixture (0.25 mM dNTP, in each case 1 µm primer 163env (SEQ ID NO:40) and primer envend (SEQ ID NO:41), 10 mM Tris HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 2.5 units of Taq polymerase (Perkin Elmer)), and amplification was then carried out in accordance with the following temperature program: 1. initial denaturation: 3 min. 95°C, 2. amplification: 90 sec. 94°C, 60 sec. 56°C, 90 sec. 72°C (30 cycles).

[082] 2. Inverse PCR

[083] The 5' region of gp 41 (N terminus) and the 3' sequence of gp 120 were amplified by means of "inverse PCR". For this, 100 µl of a genomic DNA preparation (218 µg/ml) from MVP-5180/91-infected HUT 78 cells were digested at 37°C for 1 hour in a final volume of 200 µl using 10 units of the

restriction endonuclease Sau3a. The DNA was subsequently extracted with phenol and then precipitated using sodium acetate (final concentration 300 mM) and 2.5 volumes of ethanol, with storage at -70°C for 10 min, and then centrifuged down in an Eppendorf centrifuge; the pellet was then dried and resuspended in 890 µl of distilled water. Following addition of 100 µl of ligase buffer (50 mM Tris HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg/ml bovine serum albumin) and 10 µl of T4 DNA ligase (from Boehringer, Mannheim), the DNA fragments were ligated at room temperature for 3 hours and then extracted with phenol once again and precipitated with sodium acetate and ethanol as above. After centrifuging down and drying, the DNA was resuspended in 40 µl of distilled water and digested for 1 hour with 10 units of the restriction endonuclease SacI (from Boehringer, Mannheim). 5 µl of this mixture were then employed in a PCR experiment as described under "1. PCR". The primers 168i (SEQ ID NO:42) and 169i (SEQ ID NO:43) were used for the inverse PCR in place of primers 163env (SEQ ID NO:40) and envend (SEQ ID NO:41).

[084] The primers 163env (SEQ ID NO:40), 168i (SEQ ID NO:42) and 169i (SEQ ID NO:43) were selected from that part of the sequence of the HIV isolate MVP-5180 (SEQ ID NO:56) which had already been elucidated (Example 4).

[085] The primers used for the PCR/inverse PCR and the nucleotide sequencing were synthesized on a Biosearch 8750 oligonucleotide synthesizer, with the primers having the following sequences:

Primer 163env (SEQ ID NO:40):

5' CAG AAT CAG CAA CGC CTA AAC C 3'

Primer envend (SEQ ID NO:41):

5' GCC CTG TCT TAT TCT TCT AGG 3'

(position from HIV-1 isolate BH10: bases 8129-8109)

Primer 168i (SEQ ID NO:42):

5' GCC TGC AAG CCT TAG AAA CC 3'

Primer 169i (SEQ ID NO:43):

5' GCA CTA TAC CCT TCA GTA CAC TG 3'

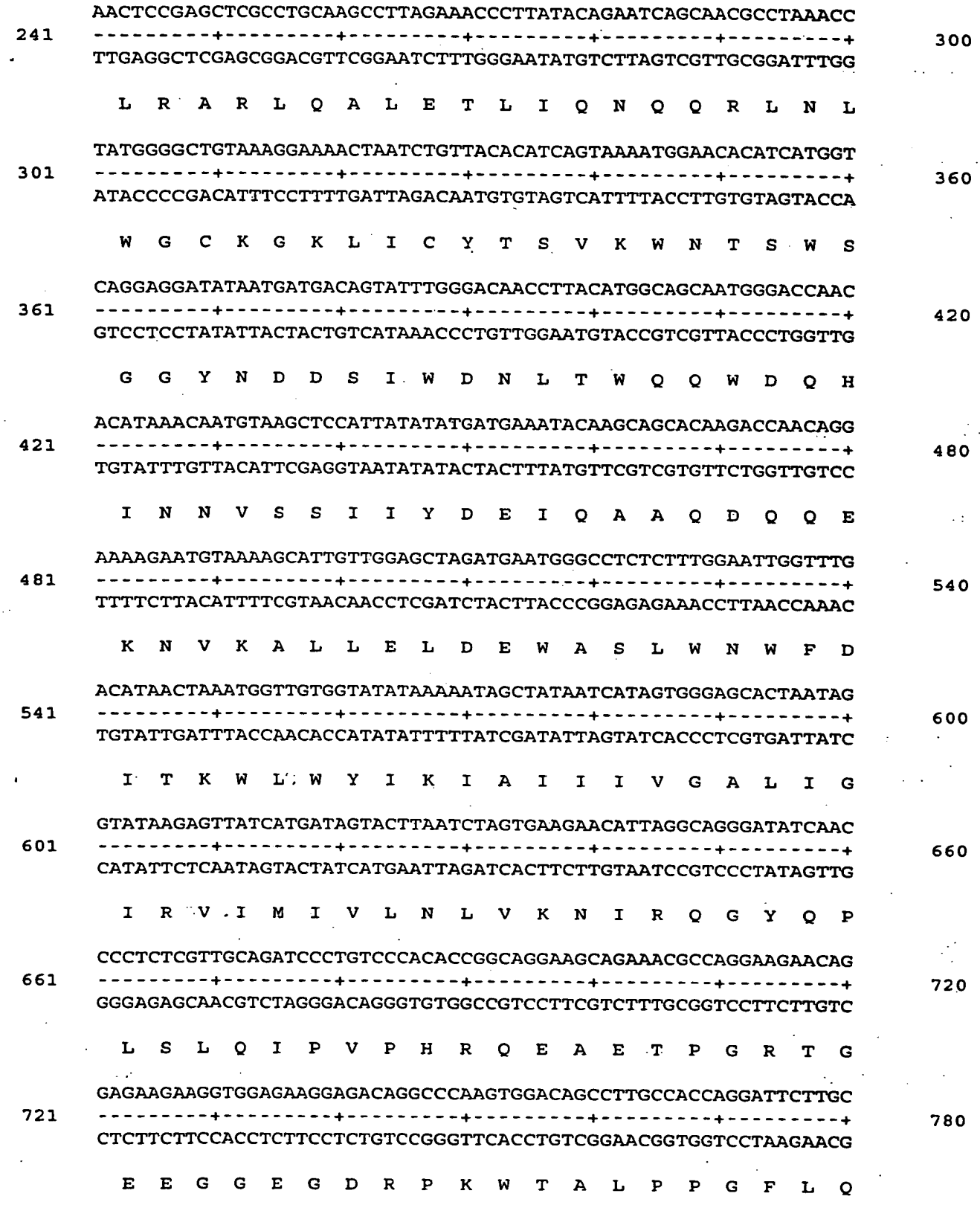
[086] The amplified DNA was fractionated on a 3% "Nusieve" agarose gel (from Biozyme) and the amplified fragment was then cut out and an equal volume of buffer (1 * TBE (0.09 M Tris borate, 0.002 M EDTA, pH 8.0)) was added to it. After incubating the DNA/agarose mixture at 70°C for 10 minutes, and subsequent phenol extraction, the DNA was precipitated from the aqueous phase by adding 1/10 vol of 3 M NaAc, pH 5.5, and 2 vol of ethanol, and storing at -20°C for 15', and then pelleted in an Eppendorf centrifuge (13,000 rpm, 10', 4°C). The pelleted DNA was dried and then taken up in water and sequenced by the method of Sanger (F. Sanger, Proc. Natl. Acad. Sci., 74: 5463, 1977) following photometric determination of the DNA concentration at 260 nm in a spectrophotometer (from Beckman). Instead of sequencing with Klenow DNA polymerase, the sequencing reaction was carried out using a kit from Applied Biosystems ("Taq dye deoxy terminator cycle sequencing", order No.: 401150). Primer 163env (SEQ ID NO:40) or primer envend (SEQ ID NO:41) (in each case 1 μM) was

employed as the primer in separate sequencing reactions. The amplified DNA from the inverse PCR experiment was sequenced using primers 168i (SEQ ID NO:42) and 169i (SEQ ID NO:43). The sequencing reaction was analysed on an Applied Biosystems 373A DNA sequencing apparatus in accordance with the instructions of the apparatus manufacturer.

[087] The nucleotide sequence of the amplified DNA region, and the amino acid sequence deduced from it, are presented in Table 3. Table 3 includes DNA sequences SEQ ID NO:44 and SEQ ID NO:45, as well as amino acid sequence SEQ ID NO:46. In Table 3, the top line corresponds to SEQ ID NO:44, the middle line corresponds to SEQ ID NO:45, and the bottom line represents amino acid sequence SEQ ID NO:46.

[088] Table 3

```
      AAATGTCAAGACCAATAATAAACATTCACACCCCTCACAGGGAAAAAAGAGCAGTAGGAT
1     ---------+---------+---------+---------+---------+---------+
      TTTACAGTTCTGGTTATTATTTGTAAGTGTGGGGAGTGTCCCTTTTTTCTCGTCATCCTA      60

        M  S  R  P  I  I  N  I  H  T  P  H  R  E  K  R |A  V  G  L
                                                 gp120      gp41

      TGGGAATGCTATTCTTGGGGGTGCTAAGTGCAGCAGGTAGCACTATGGGCGCAGCGGCAA
61    ---------+---------+---------+---------+---------+---------+
      ACCCTTACGATAAGAACCCCCACGATTCACGTCGTCCATCGTGATACCCGCGTCGCCGTT     120

        G  M  L  F  L  G  V  L  S  A  A  G  S  T  M  G  A  A  A  T

      CAGCGCTGACGGTACGGACCCACAGTGTACTGAAGGGTATAGTGCAACAGCAGGACAACC
121   ---------+---------+---------+---------+---------+---------+
      GTCGCGACTGCCATGCCTGGGTGTCACATGACTTCCCATATCACGTTGTCGTCCTGTTGG     180

        A  L  T  V  R  T  H  S  V  L  K  G  I  V  Q  Q  Q  D  N  L

      TGCTGAGAGCGATACAGGCCCAGCAACACTTGCTGAGGTTATCTGTATGGGGTATTAGAC
181   ---------+---------+---------+---------+---------+---------+
      ACGACTCTCGCTATGTCCGGGTCGTTGTGAACGACTCCAATAGACATACCCCATAATCTG     240

        L  R  A  I  Q  A  Q  Q  H  L  L  R  L  S  V  W  G  I  R  Q
```


AACTCCGAGCTCGCCTGCAAGCCTTAGAAACCCTTATACAGAATCAGCAACGCCTAAACC
241
300
TTGAGGCTCGAGCGGACGTTCGGAATCTTTGGGAATATGTCTTAGTCGTTGCGGATTTGG
L R A R L Q A L E T L I Q N Q Q R L N L
TATGGGGCTGTAAAGGAAAACTAATCTGTTACACATCAGTAAAATGGAACACATCATGGT
301
360
ATACCCCGACATTTCCTTTTGATTAGACAATGTGTAGTCATTTTACCTTGTGTAGTACCA
W G C K G K L I C Y T S V K W N T S W S
CAGGAGGATATAATGATGACAGTATTTGGGACAACCTTACATGGCAGCAATGGGACCAAC
361
420
GTCCTCCTATATTACTACTGTCATAAACCCTGTTGGAATGTACCGTCGTTACCCTGGTTG
G G Y N D D S I W D N L T W Q Q W D Q H
ACATAAACAATGTAAGCTCCATTATATATGATGAAATACAAGCAGCACAAGACCAACAGG
421
480
TGTATTTGTTACATTCGAGGTAATATATACTACTTTATGTTCGTCGTGTTCTGGTTGTCC
I N N V S S I I Y D E I Q A A Q D Q Q E
AAAAGAATGTAAAAGCATTGTTGGAGCTAGATGAATGGGCCTCTCTTTGGAATTGGTTTG
481
540
TTTTCTTACATTTTCGTAACAACCTCGATCTACTTACCCGGAGAGAAACCTTAACCAAAC
K N V K A L L E L D E W A S L W N W F D
ACATAACTAAATGGTTGTGGTATATAAAAATAGCTATAATCATAGTGGGAGCACTAATAG
541
600
TGTATTGATTTACCAACACCATATATTTTTATCGATATTAGTATCACCCTCGTGATTATC
I T K W L W Y I K I A I I I V G A L I G
GTATAAGAGTTATCATGATAGTACTTAATCTAGTGAAGAACATTAGGCAGGGATATCAAC
601
660
CATATTCTCAATAGTACTATCATGAATTAGATCACTTCTTGTAATCCGTCCCTATAGTTG
I R V I M I V L N L V K N I R Q G Y Q P
CCCTCTCGTTGCAGATCCCTGTCCCACACCGGCAGGAAGCAGAAACGCCAGGAAGAACAG
661
720
GGGAGAGCAACGTCTAGGGACAGGGTGTGGCCGTCCTTCGTCTTTGCGGTCCTTCTTGTC
L S L Q I P V P H R Q E A E T P G R T G
GAGAAGAAGGTGGAGAAGGAGACAGGCCCAAGTGGACAGCCTTGCCACCAGGATTCTTGC
721
780
CTCTTCTTCCACCTCTTCCTCTGTCCGGGTTCACCTGTCGGAACGGTGGTCCTAAGAACG
E E G G E G D R P K W T A L P P G F L Q

```
781  AACAGTTGTACACGGATCTCAGGACAATAATCTTGTGGACTTACCACCTCTTGAGCAACT
     --------+---------+---------+---------+---------+---------+
     TTGTCAACATGTGCCTAGAGTCCTGTTATTAGAACACCTGAATGGTGGAGAACTCGTTGA

       Q   L   Y   T   D   L   R   T   I   I   L   W   T   Y   H   L   L   S   N   L      840

841  TAATATCAGGGATCCGGAGGCTGATCGACTACCTGGGACTGGGACTGTGGATCCTGGGAC
     --------+---------+---------+---------+---------+---------+
     ATTATAGTCCCTAGGCCTCCGACTAGCTGATGGACCCTGACCCTGACACCTAGGACCCTG

       I   S   G   I   R   R   L   I   D   Y   L   G   L   G   L   W   I   L   G   Q      900

901  AAAAGACAATTGAAGCTTGTAGACTTTGTGGAGCTGTAATGCAATATTGGCTACAAGAAT
     --------+---------+---------+---------+---------+---------+
     TTTTCTGTTAACTTCGAACATCTGAAACACCTCGACATTACGTTATAACCGATGTTCTTA

       K   T   I   E   A   C   R   L   C   G   A   V   M   Q   Y   W   L   Q   E   L      960

961  TGAAAAATAGTGCTACAAACCTGCTTGATACTATTGCAGTGTCAGTTGCCAATTGGACTG
     --------+---------+---------+---------+---------+---------+
     ACTTTTTATCACGATGTTTGGACGAACTATGATAACGTCACAGTCAACGGTTAACCTGAC

       K   N   S   A   T   N   L   L   D   T   I   A   V   S   V   A   N   W   T   D     1020

1021 ACGGCATCATCTTAGGTCTACAAAGAATAGGACAAGG
     --------+---------+---------+-------
     TGCCGTAGTAGAATCCAGATGTTTCTTATCCTGTTCC      1057

       G   I   I   L   G   L   Q   R   I   G   Q
```

[089] **Example 8:**

[090] The found nucleotide sequence from Table 3 (SEQ ID NO:44; SEQ ID NO:45) was examined for homologous sequences in the GENEBANK database (Release 72, June 1992) using the GCG computer program (Genetic Computer Group, Inc. Wisconsin USA, version 7.1, March 1992). Most of the nucleotide sequences of immunodeficiency viruses of human origin and of isolates from primates known by July 1992 are contained in this database.

[091] The highest homology of the nucleotide sequence from Table 3 (SEQ ID NO:44; SEQ ID NO:45) to an HIV-1 isolate

is 62%. The DNA from Table 3 is 50% homologous to HIV-2 isolates.

[092] The amino acid sequence deduced from the nucleotide sequence from Table 3 (SEQ ID NO:46) was examined for homologous sequences in the SWISSPROT protein database (Release 22, June 1992) using the GCG computer program. Most of the protein sequences of immunodeficiency viruses of human origin and of isolates from primates known by June 1992 are contained in this database.

[093] At best, the amino acid sequence from Table 3 (SEQ ID NO:46) is 54% homologous to the corresponding coat protein segment from a chimpanzee isolate CIV (SIVcpz) and 54.5% homologous to the HIV-1 isolate Mal. At best, the amino acid sequence from Table 3 (SEQ ID NO:46) is 34% homologous to HIV-2 coat proteins (isolate HIV-2 D194).

[094] If, by contrast, the gp 41 amino acid sequence of HIV-1 is compared with the HIV-1 gp 41 sequence present in the SWISSPROT database, the highest homology is, as expected, almost 100%, and the lowest 78%.

[095] These clear structural differences between the sequence region from Table 3 and the corresponding segment from HIV-1 and HIV-2 suggest that isolate MVP-5180/91 (SEQ ID NO:56) is an HIV variant which clearly differs structurally from HIV-1 and HIV-2. It is possible that MVP-5180/91 (SEQ ID NO:56) should be assigned to a separate group of HIV viruses which differ from HIV-1 and HIV-2.

[096] The peptide from amino acid 584 to amino acid 618 of the HIV-1 coat protein region (SEQ ID NO:61) is of

particular serodiagnostic interest (numbering in accordance with Wain Hobson et al., Cell 40: 9-17, 1985; Gnann et al., J. Inf. Dis. 156: 261-267, 1987; Norrby et al., Nature, 329: 248-250, 1987). Corresponding amino acid regions from the coat proteins of HIV-2 and SIV are likewise immunodiagnostically conserved (Gnann et al., Science, pp. 1346-1349, 1987). Thus, peptides from this coat protein region of HIV-1 and HIV-2 are employed as solid-phase antigens in many commercially available HIV-1/2 antibody screening tests. Using them, approximately 99% of the anti-HIV-1 and anti-HIV-2-positive sera can be identified.

[097] The corresponding amino acid region of the MVP-5180/91 coat protein (Table 4), as well as the whole gp 41 of this isolate, could be of serodiagnostic importance, particularly if antisera from HIV-infected patients either did not react at all or only reacted weakly in commercially available antibody screening tests. In these cases, the infection could be due to a virus which is closely related to MVP-5180/91 (SEQ ID NO:56).

[098] Table 4 includes SEQ ID NO:61, which is designated as line 1, and also highlights in line 2 the points of difference from the amino acid sequence designated SEQ ID NO:62. Amino acid sequence SEQ ID NO:62 appears in full following Table 4.

[099] Table 4:

1 RILAVERYLKDQQLLGIWGCSGKLICTTAVPWNAS

2 LQ L TLIQN R NL K Y S K T

1 = HIV-1 amino acid sequence from gp 41 = (SEQ ID NO:61)

2 = MVP-5180 sequence from gp 41. Only differences from the HIV-1 sequence are indicated.

[100] The peptide, which was found with the aid of information deriving from MVP-5180, thus has the amino acid sequence (SEQ ID NO:62): RLQALETLIQNQQRLNLWGCKGKLICYTSVKWNTS.

[101] The present invention therefore relates to peptides which can be prepared recombinantly or synthetically and have the sequence indicated above, or a constituent sequence thereof, the constituent sequences having at least 6 consecutive amino acids, preferably 9 and particularly preferably 12 consecutive amino acids.

[102] **Example 9:**

[103] Cloning of the whole genome of the HIV isolate MVP-5180 (SEQ ID NO:56)

[104] a) Preparation of a genomic library

[105] Genomic DNA from MVP-5180-infected HUT 78 cells was isolated as described.

[106] 300 µg of this DNA were incubated for 45 min in a volume of 770 µl together with 0.24 U of the restriction enzyme Sau3A. The DNA, which was only partially cut in this incubation, was subsequently size-fractionated on a 0.7% agarose gel (low melting agarose, Nusieve) and fragments of between 10 and 21 kb were cut out. The agarose was melted at 70°C for 10 min and the same volume of buffer (1 * TBE, 0.2 M NaCl) was then added to it. Subsequently, after having extracted twice with phenol and once with chloroform, the DNA

was precipitated by adding 1/10 vol. of 3 M sodium acetate solution (pH 5.9) and 2.5 vol. of ethanol, and storing at -70°C for 10 min. The precipitated DNA was centrifuged down and dried and then dissolved in water at a concentration of 1 µg/µl.

[107] The yield of size-fractionated DNA was about 60 µg. 5 µg of this DNA were incubated at 37°C for 20 min in an appropriate buffer together with 1 U of alkaline phosphatase. In this way, the risk of multiple insertions of size-fractionated DNA was reduced by eliminating the 5'-terminal phosphate radical. The phosphatase treatment was stopped by extracting with phenol and the DNA was precipitated as above and then ligated at 15°C for 12 hours together with 1 µg of the vector (2 DASH, BamHI-cut, Stratagene No.: 247611) in a total volume of 6 µl using 2 Weiss units of Lambda T4 ligase. Following completed ligation, the DNA was packaged into phage coats using a packaging kit (Gigapack II Gold, Stratagene No.: 247611) precisely in accordance with the manufacturer's instructions.

[108] b) Radioactive labeling of the DNA probe.

[109] The "random-primed DNA labeling kit" from Boehringer Mannheim (No.: 713 023) was employed for the labeling. The PCR product was labeled which was obtained as described in Example 3 using the primers sk68 (SEQ ID NO:21) and envb (SEQ ID NO:20). 1 µg of this DNA was denatured by 2 * 5 min of boiling and subsequent cooling in ice water. 50 mCi [a- ^{32}p]-dCTP (NEN, No.: NEX-053H) were added for the

labeling. Other ingredients were added by pipette in accordance with the manufacturer's instructions. Following a 30 min incubation at 37°C, the DNA, which was now radioactively labeled, was precipitated.

[110] c) Screening the phage library

[111] 20,000 pfu (plaque-forming units) of the library in 100 µl of SM buffer (5.8 g of NaCl, 2 g of $MgSO_4$, 50 ml of 1 M Tris, pH 7.5, and 5 ml of a 2% gelatin solution, dissolved in 1 l of H_2O) were added to 200 µl of a culture (strain SRB(P2) [Stratagene, No.: 247611] in LB medium, which contained 10 mM $MgSO_4$ and 0.2% maltose) which had been grown at 30°C overnight; the phages were adsorbed to the bacteria at 37°C for 20 min and 7.5 ml of top agarose, which had been cooled to 55°C, was then mixed in and the whole sample was distributed on a pre-warmed LB agar plate of 14 cm diameter. The plaques achieved confluence after about 8 hours. After that, nitrocellulose filters were laid on the plates for a few minutes and were marked asymmetrically. After having been carefully lifted from the plates, the filters were denatured for 2 min (0.5 M NaOH, 1.5 M NaCl) and then neutralized for 5 min (0.5 M Tris, pH 8, 1.5 M NaCl). The filters were subsequently baked at 80°C for 60 min and could then be hybridized to the probe. For the prehybridization, the filters were incubated at 42°C for 2-3 h, while shaking, in 15 ml of hybridization solution (50% formamide, 0.5% SDS, 5 * SSPE, 5 * Denhardt's solution and 0.1 mg/ml salmon sperm DNA) per filter. The [^{32}P]-labeled DNA probes were denatured at 100°C for 2-5 min and then cooled on ice; they were then added

to the prehybridization solution and hybridization was carried out at 42°C for 12 hours. Subsequently, the filters were washed at 60°C, firstly with 2 * SSC/0.1% SDS and then with 0.2 * SSC/0.1% SDS. After the filters had been dried, hybridization signals were detected using the X-ray film X-OMAT™AR (Kodak).

[112] Following elution in SM buffer, those plaques to which it was possible to assign a signal were individually separated in further dilution steps.

[113] It was possible to identify the clone described below following screening of 2 * 10^6 plaques.

[114] d) Isolation of the phage DNA and subcloning

[115] An overnight culture of the host strain SRB (P2) was infected with 10 11 of a phage eluate in SM buffer such that the culture initially grew densely but then lysed after about 6-8 h. Cell remnants were separated off from the lysed culture by centrifuging it twice at 9,000 g for 10 min. Subsequently, the phage were pelleted by centrifugation (35,000 g, 1 h), and then taken up in 700 μl of 10 mM $MgSO_4$ and extracted with phenol until a protein interface could no longer be seen. The phage DNA was then precipitated and cleaved with the restriction enzyme EcoRI, and the resulting EcoRI fragments were subcloned into the vector Bluescript, KS$^-$ (Stratagene, No.: 212208). In all, 4 clones were obtained:

Plasmid	Beginning[1]	End[1]
pSP1	1	1785
pSP2	1786	5833
pSP3	5834	7415
pSP4	7660	9793

[1] refers to the total sequence below

[116] The missing section between bases 7416 and 7659 was obtained by PCR using the primers 157 (CCA TAA TAT TCA GCA GAA CTA G) (SEQ ID NO:64) and 226 (GCT GAT TCT GTA TAA GGG) (SEQ ID NO:65). The phage DNA of the clone was used as the DNA template. The conditions for the PCR were: 1.) initial denaturation: 94°C, 3 min, 2.) amplification: 1.5 min 94°C, 1 min 56°C and 1 min 72°C for 30 cycles.

[117] The DNA was sequenced as described in Example 4. Both the strand and the antistrand of the total genome were sequenced. In the case of each site for EcoRI cleavage, PCR employing phage DNA of the clone as the DNA template was used to verify that there was indeed only the one EcoRI cleavage site at each subclone transition point.

[118] Table 5: **The position of the genes for the virus proteins GAG, POL and ENV in the full sequence of MVP-5180**

Gene	Start[1]	Stop[1]
GAG	817	2310
POL	2073	5153
ENV	6260	8887

1:) The numbers give the positions of the bases in the full sequence of MVP-5180/91 (SEQ ID NO:56)

The full sequence of MVP-5180/91 is presented in Fig. 4 (SEQ ID NO:56).

[119] **Example 10:**

[120] Delimitation of the full sequence of MVP-5180/91 (SEQ ID NO:56) from other HIV-1 isolates

[121] The databanks Genbank, Release 75 of 2.93, EMBL 33 of 12.92, and Swissprot 24 of 1.93 provided the basis for the following sequence comparisons. Comparisons of homology were carried out using the GCG software (version 7.2, 10.92. from the Genetics Computer Group, Wisconsin).

[122] Initially, the sequences of GAG, POL and ENV were compared with the database at the amino acid level using the "Wordsearch" program. The 50 best homologs were in each case compared with each other using the "Pileup" program. From this, it clearly emerges that MVP-5180/91 (SEQ ID NO:56) belongs in the HIV-1 genealogical tree but branches off from it at a very early stage, even prior to the chimpanzee virus SIVcpz, and thus represents a novel HIV-1 subfamily. In order to obtain numerical values for the homologies, MVP-5180 (SEQ ID NO:56) was compared with the HIV-1, HIV-2 and SIV sequences which in each case showed the best fit, and in addition with the SIVcpz sequence, using the "Gap" program.

[123] Table 6: Homology values for the amino acid sequences of GAG, POL and ENV of the MVP-5180/91 isolate

GAG	SIVcpz	70.2% 83.6%	HIV1u[2]	69.9% 81.2%	HIV2d[3]	53.6% 71.3%	SIV1a[4]	55.1% 71.3%
POL	SIVcpz	78.0% 88.0%	HIV1u[2]	76.1% 86.8%	HIV2d[3]	57.2% 71.9%	SIVgb[5]	57.7% 74.6%
ENV	SIVcpz	53.4% 67.1%	HIV1h[1]	50.9% 67.2%	HIV2d[3]	34.4% 58.7%	SIVat[6]	34.4% 57.8%

[1]h=hz321/Zaire, [2]u=u455/Uganda, [3]d=jrcst, [4]a=agm155, [5]gb=gb1, [6]at=agm

[124] The upper numerical value expresses the identity and the lower value the similarity of the two sequences.

[125] In addition to this, the database was searched at the nucleotide level using "Wordsearch" and "Gap". The homology values for the best matches in each- case are compiled in Table 7.

[126] Table 7: Homology values for the nucleotide sequence of MVP-5180/91

	HIV1		HIV2	
gag	HIVelicg	70.24 %	HIV2bihz	60.0 %
pol	HIVmal	75.0 %	HIV2cam2	62.9 %
env	HIVsimi84	59.7 %	HIV2gha	49.8

[127] Example 11:

[128] Description of the PCR amplification, cloning and sequencing of the gag gene of the HIV 5180 isolate.

[129] In order to depict the spontaneous mutations arising during the course of virus multiplication, a part of

the viral genome was cloned using the PCR technique and the DNA sequence thus obtained was compared with the sequence according to Fig. 4 (SEQ ID NO:56).

[130] The gag sequence was cloned in an overlapping manner from the LTR (long terminal repeat, LTR1 primer) of the left end of the MVP-5180 genome through into the pol gene (polymerase gene, pol3.5i primer). The cloning strategy is depicted schematically in Fig. 5.

[131] The PCR reactions were carried out using the DNA primers given below, whose sequences were derived from the HIV-1 consensus sequence. The sequencings were carried out using the dideoxy chain termination method.

[132] The sequence encoding the MVP-5180 gag gene extends from nucleotide 817 (A of the ATG start codon) to nucleotide 2300 (A of the last codon).

LTR1 (SEQ ID NO:47): 5'-CTA GCA GTG GCG CCC GAA CAG G-3'

gag3.5 (SEQ ID NO:48): 5'-AAT GAG GAA GCU GCA GAU TGG GA-3'

(U=A/T)

gag 3.5i (SEQ ID NO:49): 5'-TCC CAU TCT GCU GCT TCC TCA TT-3'

(U=A/T)

gag5 (SEQ ID NO:50): 5'-CCA AGG GGA AGT GAC ATA GCA GGA AC-3'

gag959 (SEQ ID NO:51): 5'-CGT TGT TCA GAA TTC AAA CCC-3'

gag11i (SEQ ID NO:52): 5'-TCC CTA AAA AAT TAG CCT GTC-3'

pol3.5i (SEQ ID NO:53): 5'-AAA CCT CCA ATT CCC CCT A-3'

[133] The DNA sequence obtained using the PCR technique was compared with the DNA sequence presented in Figure 4 (SEQ

ID NO:56). A comparison of the two DNA sequences is presented in Figure 6. Figure 6 includes SEQ ID NO:57, which corresponds to Fig. 4 (SEQ ID NO:56) and SEQ ID NO:58, which corresponds to the DNA sequence obtained using the PCR technique. This showed that about 2% of the nucleotides differ from each other, although the virus is the same in the two cases. In Fig. 6, the upper line in each case represents the DNA sequence which is presented in Fig. 4 (SEQ ID NO:56) and the lower line represents the DNA sequence obtained using the PCR technique.

[134] In addition, the amino acid sequence of the gag protein, elucidated using the PCR technique, was compared with the amino acid sequence of the corresponding protein deduced from Fig. 4 (SEQ ID NO:59). This showed an amino acid difference of about 2.2%. The comparison is presented in Fig. 7, the lower line in each case representing the amino acid sequence which was deduced from the sequence obtained using the PCR technique. Figure 7 includes amino acid SEQ ID NO:59, which was elucidated in accordance with Fig. 4 (SEQ ID NO:56), and the amino acid sequence (SEQ ID NO:60) derived using the PCR technique.

[135] **Example 12:**

[136] The sequence of the virus MVP-5180 (SEQ ID NO:56) according to the invention was compared with the consensus sequences of HIV-1 and HIV-2, and with the sequence of ANT-70 (WO 89/12094), insofar as this was known.

[137] In this connection, the following results were obtained:

[138] **Table 8:**

Gene locus	Deviating nucleotides	Number of the nucleotides	% homology (approximated)	
LTR	207 308 115	630	HIV-1 HIV-2 ANT 70	67 % 51 % 82 %
GAG	448 570	1501	HIV-1 HIV-2	70 % 62 %
POL	763 1011	3010	HIV-1 HIV-2	74% 66%
VIF	183 338	578	HIV-1 HIV-2	68% 42%
ENV	1196 1289	2534	HIV-1 HIV-2	53% 49%
NEF	285 342	621	HIV-1 HIV-2	54% 45%
total	3082 3858	8874	HIV-1 HIV-2	65% 56%

[139] In the above table, "HIV-1" denotes consensus sequences of HIV-1 viruses; "HIV-2" denotes consensus sequences of HIV-2 viruses; ANT-70 denotes the partial sequence of a virus designated HIV-3 and disclosed in WO 89/12094.

[140] The present invention therefore relates to viruses, DNA sequences and amino acid sequences, and constituent sequences thereof, which possess such a degree of homology with the sequence presented in Fig. 4 (SEQ ID NO:56), based on the gene loci, that at most the fractions given in Table 9, expressed in % values, are different.

[141] Table 9: Homology based on gene loci, expressed as maximum differences

Gene locus	Differences	Preferred differences	Particularly preferred differences
LTR	17 %	15 %	10 %
GAG	29 %	28 %	14 %
POL	25 %	24 %	12 %
VIF	31 %	30 %	15 %
ENV	46 %	45 %	22 %
NEF	16 %	12 %	10 %

[142] The homology values in % given in Table 9 mean that, when comparing the sequence according to Fig. 4 (SEQ ID NO:56) with a sequence of another virus, at most a fraction of the sequence corresponding to the abovementioned percentage values may be different.

[143] **Example 13:** V3 loop

[144] This loop is the main neutralizing region in HIV and the immunological specificities of the region are documented in summary form in Figure 8. This is a copy from a work by Peter Nara (1990) from AIDS. The amino acid sequence of the V3 loop is shown diagrammatically and is compared with the IIIB virus, now LAI, and the first HIV-2 isolate (ROD). Individual amino acids are conserved at the cystine bridge. Whereas the crown of HIV-1 is GPGR or GPGQ and that of HIV-2 is GHVF, the crown of MVP-5180/91 (SEQ ID NO:56) is formed from the amino acids GPMR. The motif with methionine has not previously been described and, emphasizes the individuality of MVP-5180/91 (SEQ ID NO:56).

[145] After having determined the nucleotide sequence of the virus the V3-loop-region was amplified using the PCR-technique by using suitable primers. Some mutations have been observed, especially a change of the methionine codon (ATG) to the leucine codon (CTG).

[146] In the following the amino acid sequence derived from the cloned nucleic acid is compared with a sequence obtained after amplification with the help of PCR technology.

MvP 5180 (cloned) (SEQ ID NO:54):
CIREGIAEVQDIYTGPMRWRSMTLKRSNNTSPRSRVAYC

MvP 5180 (PCT technique) (SEQ ID NO:55):

CIREGIAEVQDLHTGPLRWRSMTLKKSSNSHTQPRSKVAYC

[147] **Example 14:**

[148] In order to demonstrate that even those sera which cannot be identified in a normal HIV-1+2 screening test can be proved to be HIV-1-positive with the aid of the virus MVP-5180 (SEQ ID NO:56) according to the invention, or antigens derived therefrom, various sera from patients from the Cameroons were examined in the EIA test.

[149] 156 anti-HIV-1-positive sera were examined in a study carried out in the Cameroons. Substantial, diagnostically relevant differences were observed in two of these sera. The extinctions which were measured are given in

Table 10 below. CAM-A and CAM-B denote the sera of different patients.

[150] Table 10:

Patient sera	MVP-5180-EIA	HIV-1 + HIV-2 EIA
CAM-A	2.886	1.623
CAM-B	1.102	0.386

The cutoff for both tests was 0.300.

[151] In a further study on 47 anti-HIV-1-positive sera from the Cameroons, two sera were of particular note. One of these (93-1000) derives from a patient showing relatively few symptoms and the other (93-1001) from a patient suffering from AIDS. The extinction values for the two EIA tests are compared in Table 11 below:

[152] Table 11:

Patient sera	MVP-5180-EIA	HIV-1 + HIV-2 EIA
93-1000	> 2.5	1.495
93-1001	0.692	0.314

[153] The cutoff was 0.3 in this case as well. The extinction values for patient 93-1001 demonstrate that the normal HIV-1 + HIV-2 EIA can fail whereas clear detection is possible the antigen according to the invention is employed.